Deep learning based depth map estimation of protoporphyrin IX in turbid media using dual wavelength excitation fluorescence

HINANO IMANISHI,[1] TAKAHIRO NISHIMURA,[1,*] YU SHIMOJO,[1,2,3] (iD) AND KUNIO AWAZU[1,4]

[1]*Graduate School of Engineering, Osaka University, Yamadaoka 2-1, Suita, Osaka 565-0871, Japan*
[2]*Graduate School of Medicine, Osaka Metropolitan University, Asahimachi 1-4-3, Abeno-ku, Osaka 545-8585, Japan*
[3]*Research Fellow of Japan Society for the Promotion of Science, Kojimachi 5-3-1, Chiyoda-ku, Tokyo 102-0083, Japan*
[4]*Global Center for Medical Engineering and Informatics, Osaka University, Yamadaoka 2-2, Suita, Osaka 565-0871, Japan*
*nishimura-t@see.eng.osaka-u.ac.jp

Abstract: This study presents a depth map estimation of fluorescent objects in turbid media, such as biological tissue based on fluorescence observation by two-wavelength excitation and deep learning-based processing. A U-Net-based convolutional neural network is adapted for fluorophore depth maps from the ratiometric information of the two-wavelength excitation fluorescence. The proposed method offers depth map estimation from wide-field fluorescence images with rapid processing. The feasibility of the proposed method was demonstrated experimentally by estimating the depth map of protoporphyrin IX, a recognized cancer biomarker, at different depths within an optical phantom.

1. Introduction

Fluorescence imaging has improved tumor diagnosis by providing real-time and high-contrast capabilities [1,2]. Photodynamic diagnosis (PDD) using 5-aminolevulinic acid (ALA) has been widely employed for intraoperative tumor diagnosis in various areas, including malignant glioma [3], bladder cancer [4], and prostate cancer [5]. ALA administration leads to protoporphyrin IX (PpIX) accumulation, specifically in tumor tissue [6]. The fluorescence emission from PpIX enables high-contrast visualization of tumor areas. The selective accumulation of PpIX has been observed in various tumor types, providing its application in the treatment of different cancers [7]. In addition, it can improve imaging in PpIX-PDD, such as quantifying tumor depth in tissue [8–10]. Accurate quantification of tumor depth within tissue enables surgical planning and treatment strategy determination. Information on tumor depth assists surgeons in managing margins more appropriately within the resection area [11]. This could lead to shorter surgery durations and reduce the likelihood of requiring follow-up operations. Information on tumor depth also aids in the diagnosis of tumor invasion. Endoscopic resection for early gastric cancer is carried out for patients with tumors situated within the gastric mucosa or the submucosal layer [12,13]. It is anticipated that intraoperative tumor depth diagnosis using endoscopy will be utilized for patient screening prior to endoscopic resection.

A method based on the linear relationship between fluorescence intensity ratio and fluorophore depth has been developed to measure the depth of fluorophores within biological tissue [8–10]. This technique leverages the wavelength-dependent optical properties of the tissue. By acquiring fluorescence images at different excitation wavelengths, variations in fluorescence intensity arise owing to the wavelength-dependent attenuation of the excitation light [14]. The relationship

between fluorescence intensity ratio and fluorophore depth can be modeled using the attenuation of excitation light at each wavelength in advance. Based on the relationship between the fluorescence intensity ratio and fluorophore depth, the fluorophore depth information can be estimated from the measured fluorescence intensity ratio. This approach is promising for clinical applications, such as intraoperative use, as it utilizes wide-field optical systems that allow for relatively simple data acquisition and processing [15]. However, when multiple fluorescent objects are present with different positions and depths, the relationship between fluorescence intensity ratio and fluorophore depth does not hold, reducing accuracy [10,16]. A fluorescence depth estimation method is needed for several locations using fluorescence intensity ratios.

Here, we present a method employing a deep learning model to an estimation depth map of multiple fluorescent objects consisting of PpIX from images of two-wavelength excitation fluorescence using a measured fluorescence intensity ratio map. To estimate a depth distribution of PpIX within a turbid media, a fluorescence ratio map from two fluorescence images acquired through two-wavelength excitation is prepared. The fluorescence ratio map image is input into a deep learning model, which utilizes a convolutional neural network (CNN) architecture to output the depth map of PpIX. The CNN model is trained using a dataset of fluorescence images generated through Monte Carlo simulations. Utilizing the deep learning model allows our proposed method to handle cases where multiple fluorescent objects exist, and estimation accuracy is decreased with a linear model of the fluorescence ratio. The utilization of wide-field imaging and real-time processing holds the promise of enabling the rapid acquisition of depth information, ultimately contributing to advancements in fluorescence tumor diagnosis.

2. Materials and methods

2.1. PpIX depth map prediction using U-net

Figure 1 shows the procedure for depth map estimation using a deep learning model. Fluorescence images, $F_b(x, y)$ and $F_g(x, y)$, are obtained with excitation at wavelengths of 405 and 505 nm. The excitation wavelengths of 405 and 505 nm were selected from the absorption peak wavelengths in the region below 600 nm where the fluorescence wavelength band of PpIX does not overlap [14]. The 405-nm wavelength light irradiation excites around the maximum absorption peak and is attenuated by tissue absorbance and scattering. This leads to strong excitation of the PpIX distributed near the tissue surface and little excitation of the PpIX distributed deeply in tissues. In contrast, the 505-nm wavelength excitation light, which excites smaller absorption peaks than 405 nm, can be delivered to the PpIX distributed deeply in tissues. For example, the optical penetration depths where the light intensity decreased to $1/e$ of porcine gastric mucosal tissue were reported as 0.26 mm and 0.66 mm for 405 nm and 505 nm of light excitation, respectively [14]. The 505-nm wavelength excitation light can increase the fluorescence intensity of PpIX distributed deeply in tissues compared to 405-nm light irradiation. A fluorescence ratio map $\Gamma(x, y)$ is defined as below:

$$\Gamma(x, y) = \begin{cases} \frac{F_b(x,y)+F_g(x,y)}{F_g(x,y)} & \text{if } F_g(x, y) \neq 0, \\ 1 & \text{if } F_g(x, y) = 0. \end{cases} \tag{1}$$

The fluorescence ratio map is fed into a neural network with a U-Net structure to obtain a depth map of PpIX objects [17]. The neural network is trained to output a depth map of the PpIX objects. The resulting depth map does not distinguish whether PpIX was present or deeply located, as $\Gamma(x, y)$ indicated 1 for both conditions. To address this, the region that indicated the estimated depth of the area where PpIX fluorescence cannot be detected by 505-nm excitation is judged as an area where PpIX is absent. Thus, a PpIX depth map can be obtained.

Figure 2 shows the U-Net-based architecture to estimate a fluorescence depth map from the PpIX fluorescence images obtained by 405- and 505-nm light irradiation. The model
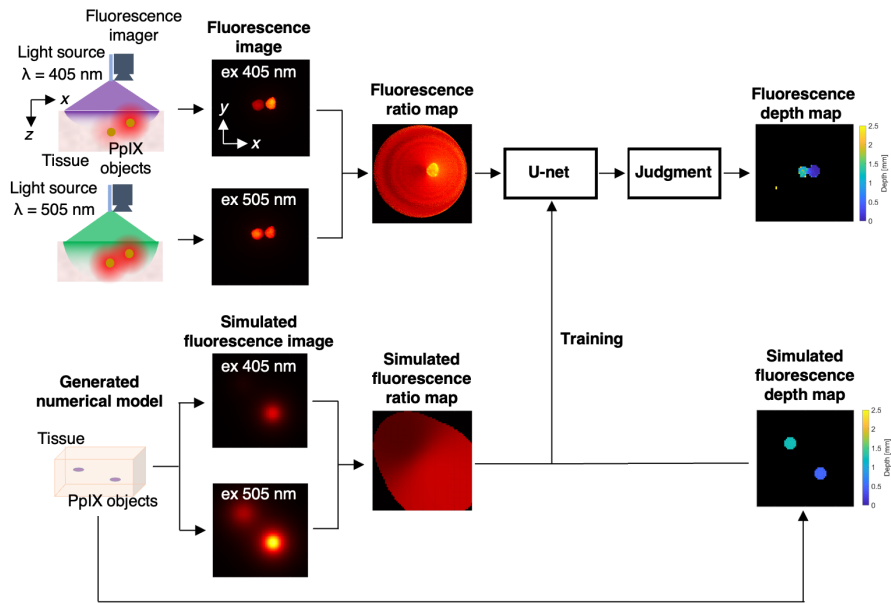
Fig. 1. Schematic of depth map estimation from fluorescence ratio map of dual-excited PpIX fluorescence images.

comprises 29 layers, including input, output, and convolutional layers. The number of filters in the convolutional layers was set to 64, 128, 256, 512, and 1024, with a kernel size of 3×3. A regression model was designed to output the optimal depth map for the input fluorescence ratio image. The network adopted the ReLU activation function [18]. The ReLU activation function, widely used in neural networks, outputs zero for any negative input and retains positive input as is. The Adam optimization algorithm [19] was employed. The Adam algorithm is a method in machine learning for solving optimization problems in complex systems. Root mean squared error between the estimated and the ground-truth depth maps was employed as a loss function. The network was trained using the hold-out method. Only numerically simulated data was used for training and validation. For training and validation, simulated datasets of pairs of the fluorescence ratio image and the depth map were prepared. The dataset was split into an 80:20 ratio, allocating 80% for training and 20% for validation. The network was trained with epochs = 40, batch size = 475, and learning rate = 0.0001.
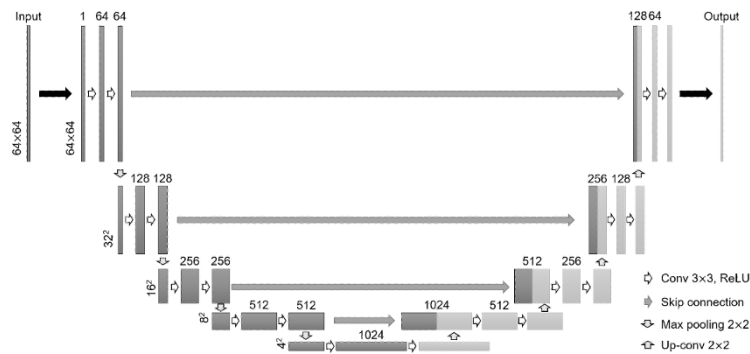

Fig. 2. Adopted U-Net architecture for depth map estimation.

2.2. Simulated fluorescence image

For datasets, $F_b(x, y)$ and $F_g(x, y)$ were generated computationally. $F_b(x, y)$ and $F_g(x, y)$ were assumed as the PpIX fluorescence intensity distributions on the tissue surface. $F_b(x, y)$ or $F_g(x, y)$ can be expressed as follows.

$$F_b(x, y) = \int \int \int \phi(x - x', y - y', -z') P_b(x', y', z') \varepsilon_b O(x', y', z') Q dx' dy' dz', \qquad (2)$$

$$F_g(x, y) = \int \int \int \phi(x - x', y - y', -z') P_g(x', y', z') \varepsilon_g O(x', y', z') Q dx' dy' dz', \qquad (3)$$

where $\phi(x, y, z)$, $P_b(x', y', z')$, $P_g(x', y', z')$, ε_b, ε_g, $O(x', y', z')$ and, Q are a point spread function for PpIX fluorescent imaging, excitation distribution for 405 and 505 nm, molar absorption coefficients for 405 and 505 nm, PpIX concentration, and quantum yield of PpIX. In this study, for simplicity, the PpIX concentration was represented by binary values of 0 and 1, indicating the absence or presence of PpIX, respectively. Also, $F_b(x, y)$ and $F_g(x, y)$ were normalized by the fluorescence intensities of the PpIX at the surface ($z = 0$), to cancel factors, ε_b, ε_g, and Q in Eqs. (2) and (3). The normalized $F_b(x, y)$ and $F_g(x, y)$ were prepared and used for training and evaluation of the U-Net model. The functions of $P_b(x', y', z')$, $P_g(x', y', z')$ were calculated by a Monte Carlo (MC)-based calculation of light transportation in an optical phantom simulating tissue. The MC-based calculations were performed using a software package named Monte Carlo eXtreme (MCX) [20]. The numerical optical model comprised a cube with a size of $400 \times 400 \times 400$ voxels and a voxel size of $0.25 \times 0.25 \times 0.25$ mm. The upper 20 mm of the model was set as an air layer and the rest as a tissue layer. The optical properties adopted in the simulation are presented in Table 1. The optical properties of the phantom were set to values of an optical phantom employed in experimental verification. Anisotropy factors were assumed as 1 for the air layer and 0.9 for the phantom layer. The number of photons for the MC calculation was set to 10^8. The excitation lights at 405 and 505 nm were assumed to be a circular light source with a diameter of 100 mm and were irradiated perpendicular to the tissue layer. $\phi(x, y, z)$ is defined as the fluorescence spread at the surface of the tissue from the PpIX point source at a coordinate (x, y, z). $\phi(x, y, z)$ was calculated using MCX with an interval of 0.1 mm between z = 0–2.5 mm in advance. The wavelength of the PpIX fluorescence was approximated to the single wavelength of 635 nm. The center 64×64 pixels (16×16 mm) of the calculated fluorescence intensity distribution was cropped and adopted as a simulated fluorescence image.

Table 1. Absorption and scattering coefficients (μ_a and μ_s) employed in simulation.

	Air layer 405, 505, 635 nm	Phantom layer 405 nm	Phantom layer 505 nm	Phantom layer 635 nm
μ_a (mm^{-1})	1.0×10^{-3}	0.65	0.10	0.04
μ_s (mm^{-1})	0.1	17.3	9.1	5.2

2.3. Dataset generation

Numerical tissue models, including two PpIX pellets, were generated for the datasets. The shape of PpIX pellets was circular disks with their thickness and diameter adjusted according to the evaluation contents. Two types of datasets, Datasets A and B, were prepared. Dataset A comprises 8,239 pairs of fluorescence ratio and depth maps, where the diameter of PpIX pellet was 2 mm. Dataset B comprises 24,837 pairs of fluorescence ratio and depth maps, where the diameters of PpIX pellet were 1, 2, or 3 mm. The PpIX pellets were arranged horizontally on the surface of the tissue. The thickness of the PpIX pellet was ignored (= 0 mm) in the simulations. The fluorescence depth was defined as the length between the tissue surface and the PpIX pellets.

The two PpIX pellets were randomly placed within the fluorescence depth of 2.5 mm with 0.1 mm resolution. Fluorescence images by excitation of 405 and 505 nm were simulated for each numerical model, as shown in Sec.2.2. The fluorescence ratio maps were calculated from the sets of fluorescence images by Eq. (1). The fluorescence depth map was also created from the location information of the PpIX pellets in the numerical tissue model. The depth was defined as the deeper value in the overlapped region of the PpIX pellets in the depth map. Thus, the fluorescence ratio map sets and the corresponding depth map were prepared. Each dataset was split into an 80:20 ratio, allocating 80% for training and 20% for validation.

For test dataset, fluorescent images with two PpIX pellets were placed at varying depths. Ten data for each of the three scenarios: the positions of two PpIX pellets on the depth map were separated, touched, and overlapped. The positions of the PpIX pellets in the *x* and *y* directions were manually set, while the other procedures were the same as those in the training dataset. 10 data were prepared for each scenario.

2.4. Evaluation index

Two indices were utilized to assess the performance of the trained network. The Dice coefficient was adopted to evaluate the estimated location of PpIX fluorescence in the *xy*-direction. The Dice coefficient, D, is denoted as below:

$$D = \frac{2 \times N_{E \cap T}}{N_E + N_T},\tag{4}$$

where N_E and N_T represent the number of pixels in the regions where the PpIX is assumed to exist in the estimated and ground truth depth maps, respectively. $N_{E \cap T}$ represents the number of pixels where the PpIX is assumed to exist in both the ground truth and estimated depth maps. Estimation error was calculated as the difference between the estimated depth map and the ground true depth map. For quantitative comparison, mean absolute error (MAE) was calculated within the region where fluorescence is assumed to exist in both the estimated and ground truth depth maps.

2.5. Optical setup

The optical setup utilized for fluorescence observation is illustrated in Fig. 3(a). Fiber-coupled laser sources with wavelengths of 405 (WSLS-405-002-M, Wavespectrum Laser.Inc.) and 505 nm (WSLS-505-020m-PM, Wavespectrum Laser.Inc.) were employed as excitation light sources. The light sources were switched for the fluorescence images for the 405- and 505-nm light excitation. The fiber ends were fixed and adjusted to make the excitation light almost uniform on the sample surface within the observation field. The excitation light intensities on the sample surface were measured using a power sensor (PD300-UV, OPHIR) with a display (NOVA II, OPHIR) attached to the sensor head. The excitation light intensity was set as 12.8 mW/cm^2 at 405 nm and 1.27 mW/cm^2 at 505 nm on the sample plane. Fluorescence images were captured using a CMOS monochrome camera (Chameleon3, CM3-U3-31S4M, FLIR) and a lens (f=16 mm, F1.4, NMV-16M1, Navitar). A long-pass filter (FELH0600, Thorlabs) was placed in front of the camera lens to cut off the excitation light.

2.6. Sample preparation

Optical phantoms and PpIX pellets were prepared in the same manner as the previous reports [14,21]. Hemoglobin, TiO$_2$, and gelatin were used as the absorber, scatterer, and support, respectively. The final concentrations of hemoglobin, TiO$_2$, and gelatin were adjusted to 0.5 mg/ml, 1.0 mg/ml, and 50 mg/ml. The absorbance and scattering coefficients of the optical phantom were measured as 0.65 mm^{-1} ± 0.04 and 17.3 mm^{-1} ± 0.7 for a wavelength of 405 nm, 0.10 mm^{-1} ± 0.01 and 9.1 mm^{-1} ± 0.3 for 505 nm, and 0.04 mm^{-1} ± 0.01 and 5.2 mm^{-1} ± 0.2
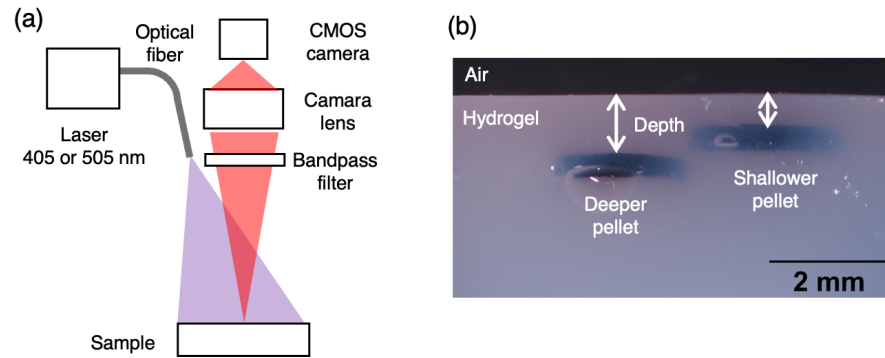
Fig. 3. (a) Optical setup. (b) Cross section of phantom.

for a wavelength of 405 nm, using a double-integrating sphere in the same manner as described in the Ref. [22]. The diameter of the PpIX pellet was 2 mm. The PpIX pellets were located in the optical phantom at a depth of around 0.5 and 1 mm. To measure the depth of PpIX pellets, after fluorescence image acquisitions, the optical phantom was sliced, and the cross-section was observed by a microscope to measure the depth of the PpIX pellets (Fig. 3(b)). Six samples were prepared in total for experimental evaluation.

2.7. Fluorescence image acquisition

Fluorescence images of a sample were obtained by excitation light wavelengths of 405 and 505 nm. Each sample was imaged three times per excitation wavelength, and the averages were taken as measured fluorescence images. The exposure times were adjusted in the range of 1–70 ms for each sample to prevent saturation of the pixel values in the measured fluorescence image. Then, the pixel values of all fluorescence images were adjusted to match the pixel values of the fluorescence image taken at 70 ms of exposure time, based on the linearity between the exposure time and brightness [14]. As calibration fluorescence images, fluorescence images of a sample where the PpIX pellet was located on the phantom surface were taken for each excitation light wavelength. The pixel values of all fluorescence images were normalized to the maximum brightness of the calibration fluorescence images per each excitation wavelength. To input the acquired fluorescence intensity ratio images into a trained deep learning model, the central areas were trimmed and resized to 64×64 pixels using a Gaussian filter to match the size and resolution of the simulated fluorescence images.

3. Results

3.1. Numerical verification

To demonstrate depth map estimation, the network was trained using Dataset A. The training time was 23.7 min on a computer with an NVIDIA GeForce RTX 3060 Laptop GPU. The average calculation time was 21 ms to estimate a fluorescence depth map. Figure 4 shows an example of estimation results from simulated fluorescence images for three conditions: separated, touched, and overlapped. The Dice coefficients and MAEs are listed in Table 2. The Dice coefficients slightly decreased for the touched and overlapped cases compared to the separated ones. In the worst case, the average and 2σ of MAE were 0.29 and 0.18 mm, which means that 95% of the absolute error was suppressed within 0.47 mm.
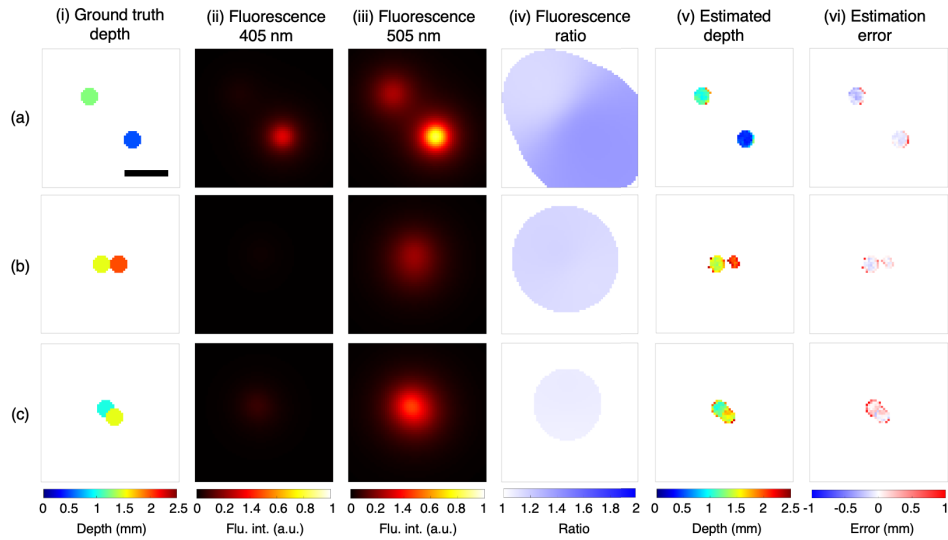
Fig. 4. Numerical evaluation results from simulated fluorescence images for three conditions: (a) separated, (b) touched, and (c) overlapped pellets. (i) Ground truth depth map. Simulated fluorescence images by (ii) 405-nm (iii) 505-nm excitation light. (iv) Fluorescence ratio map. (v) Estimated map. (vi) Estimation error map. Black scalebar indicates 5 mm.

Table 2. Quantitative analysis of numerical experiments ($n = 10$ for each scenario).

Sample type	Dice coefficient	Mean absolute error [mm]
Separated	0.95 ± 0.03	0.16 ± 0.05
Touched	0.88 ± 0.08	0.18 ± 0.03
Overlapped	0.88 ± 0.07	0.29 ± 0.09

3.2. Numerical evaluation of robustness to tissue optical property

To evaluate the effect on the estimation accuracy by the variances of the optical property values, depth estimation was performed by inputting fluorescence images generated by varying the absorption and scattering coefficients of the optical phantom layer by ± 5 %, ± 10 %, and ± 15 %. The values for the absorption and scattering coefficients are provided in Supplemental Table S1 for further reference. The ± 15% variations in optical properties were set based on the observed variability within the optical phantoms we experimentally fabricated. When altering the optical properties, absorption and scattering coefficients increased or decreased together. The other conditions and the trained network were the same as the verifications in Sec. 3.1. Figures 5(a,b) show estimation results from simulated fluorescence images for the same condition with the separated location shown in Fig. 4(a) but the optical property values were varied ± 15 %. For the case of -15 % optical property values (Fig. 5(a)), the mean estimated error in the shallower pellet was -0.13 mm, and that in the deeper pellet was -0.14 mm. Both pellets were estimated to be shallower in the center. In the case of a $+15$ % optical property value (Fig. 5(b)), the mean estimated error in the shallower pellet $+0.16$ mm and that in the deeper pellet was $+0.12$ mm. The PpIX pellets were estimated to be deeper in both conditions. Because the absorption and scattering coefficients of the optical phantom at a wavelength of 405 nm are larger than those at a wavelength of 505 nm, the influence of changing the optical property values is significant. Increased optical properties at 505 nm reduce the fluorescence intensity ratio due to the greater

optical penetration depth of the 505 nm excitation light, leading to estimations of a shallower location. In contrast, decreased optical properties at 505 nm increase the fluorescence intensity ratio, leading to estimations of a deeper position. Figures 5(c, d) show D and MAE in-depth estimation of the pellets in phantoms with varied optical properties. The MAE increased with the gaps in the optical properties between the simulated model for generating the training dataset and the actual sample.
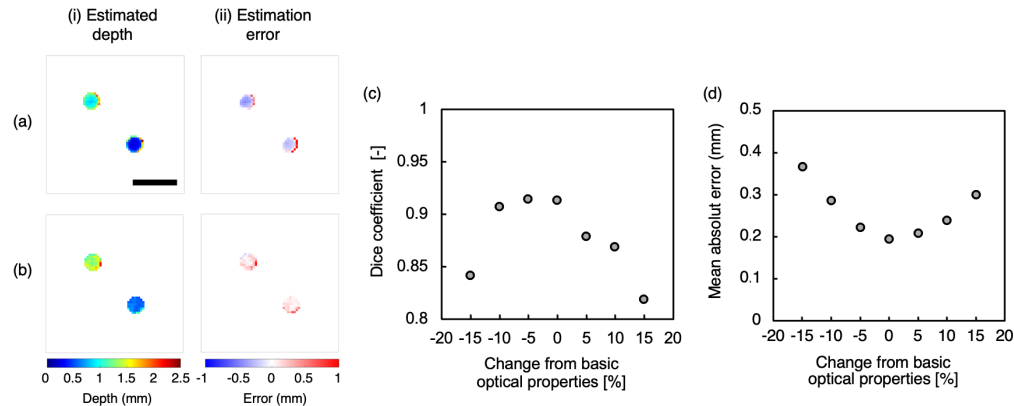

Fig. 5. Examples of depth estimation results simulated fluorescence images of PpIX in phantoms with (a) −15 %- and (b) +15 %-varied optical property values. Relationship between change of optical properties and evaluation items: (c) Dice coefficient and (d) mean absolute error. Black scalebar indicates 5 mm

3.3. Numerical evaluation of robustness to fluorescence structure

To evaluate the robustness of the PpIX distributions, we compared models trained on Datasets A and B. Using each model, we estimated depth maps for two disk-shaped pellets with diameters of either 1 or 3 mm. An example of the ground-truth depth map, estimated depth map, and comparison map for each condition in depth estimation using the simulated fluorescence images is shown in Figs. 6. By training with Dataset A, D and MAE were 0.57 and 0.99 mm for a diameter of 1 mm (Fig. 6(a)), and D and MAE were 0.56 and 0.70 mm for a diameter of 3 mm (Fig. 6(b)). In Fig. 6(a.ii), two objects with a diameter of about 2 mm were observed at the true position of the PpIX pellet with a diameter of 3 mm. In Fig. 6(b.ii), two objects with a diameter of about 2 mm patterns were observed around the position where the PpIX pellet with a diameter of 1 mm should be. These indicate that the trained model might lose its generalization performance for the PpIX distributions. By training with Dataset B, D was 1.0, and MAE was 0.20 mm for a diameter of 1 mm, while D was 0.97 and MAE was 0.09 mm for a diameter of 3 mm. Furthermore, Fig. 7 shows the correct depth map, estimated depth map, and comparison map for a 1.4- or 2.4-mm diameter not included in the training dataset and the estimated depth map generated by the new trained model. For a diameter of 1.4 mm, D and MAE were 0.72 and 0.28 mm. For a diameter of 2.4 mm, D and MAE were 0.93 and 0.17 mm. These results indicate that increasing the variation in the training dataset improves the estimation accuracy for unknown shapes. In clinical use, it is suggested that reflecting the characteristics of the diagnostic target, such as tumor shape, in the fluorescent intensity ratio image of the training dataset can lead to more accurate depth map estimation. To accurately reflect the structural characteristics of actual tumors, it is required to expand the field of view in the fluorescent imaging data used for training. This need arises from the fact that tumor sizes usually exceed the size of both simulated and experimental fluorescence images employed in this study.
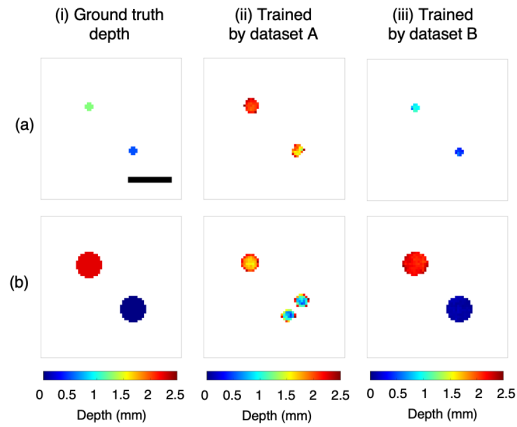
Fig. 6. Depth estimation results of fluorophores of (a) 1 mm and (b) 3 mm diameter pellets from simulated fluorescence images. (i) Ground truth depth map. Estimated depth map by the neural networks trained by (ii) Dataset A and (iii) Dataset B. Black scalebar indicates 5 mm.
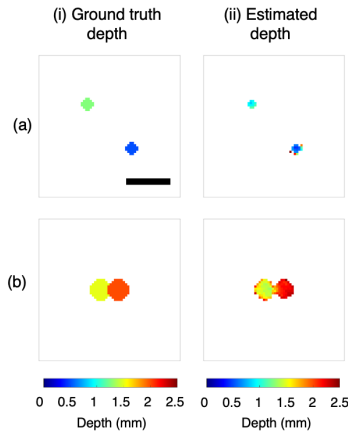

Fig. 7. Depth estimation results of fluorophores of (a) 1.4 mm and (b) 2.4 mm diameter pellets from simulated fluorescence images. (i) Ground truth depth map. Estimated depth map by neural networks trained by (ii) Dataset B. Black scalebar indicates 5 mm.

3.4. Experimental validation of fluorescence depth estimation

Fluorescence ratio maps were obtained from experimentally acquired fluorescence images, and the depth maps were estimated using the network trained by Dataset A. Figure 8 shows the estimation results of depth maps from the fluorescence images acquired experimentally with the optical setup. The results of estimating six samples were D of 0.73 ± 0.08 and MAE of 0.46 ± 0.11 mm. Figure 9 shows the measured and estimated pellet depths for all samples. MAE was 0.74 ± 0.35 mm for shallower pellets and 0.22 ± 0.16 mm for deeper pellets, indicating an overall tendency to estimate deeper.
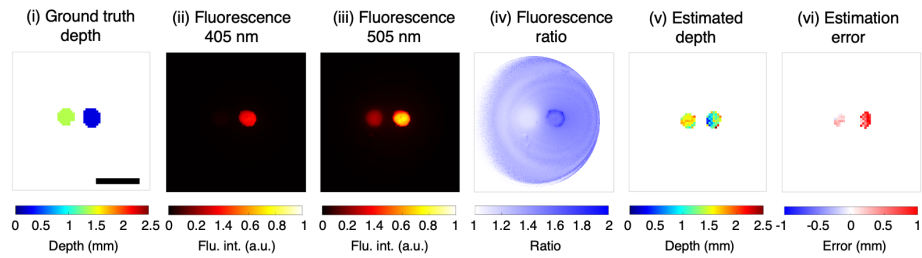
Fig. 8. An experimental result from fluorescence images acquired experimentally.(i) Ground truth depth map. Measured fluorescence images by (ii) 405-nm (iii) 505-nm excitation light. (iv) Fluorescence ratio map. (v) Estimated map. (vi) Estimation error map. Black scalebar indicates 5 mm.
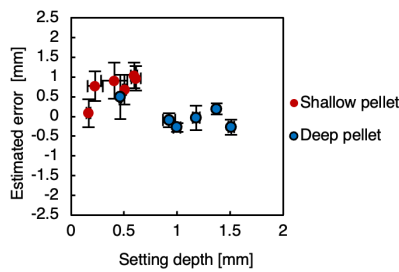

Fig. 9. Relationship between setting depth and estimated error of fluorescence object in depth estimation from fluorescence images acquired experimentally. Error bar in *x* axis indicates standard deviation of setting depth. Error bar in *y* axis indicates standard deviation of estimated depth.

4.   Discussions

Numerical and experimental results demonstrated that the U-Net-based convolution network could estimate the depth map of the PpIX sheeted in tissues by using the fluorescence intensity ratio information. From the fluorescence ratio map, which contains the fluorescence information from different depths, the depth of each PpIX pellet can be estimated with simple and fast processing ($\simeq$ 21 ms). The numerical validation revealed the accurate estimation result with a Dice coefficient of 0.91 and an absolute error of 0.19 mm. In experiments, the depth map can be provided only by obtaining dual excitation fluorescence only by fluorescence excitation wavelength switching. Due to the effects of attenuation by the tissue of both excitation light and fluorescence, the fluorescence of PpIX near the superficial layer is dominant in fluorescence images. However, as shown in Fig. 9, the experimental results yielded a smaller estimated error in the estimation of pellets located deeper within the tissue. Although this observation may be influenced by experimental variables, such as pellet thickness, it indicates that machine learning is capable of effectively estimating depths for multiple targets, including those situated deeper within the tissue. Within the depth range of a few millimeters considered in this study, our findings suggest that three-dimensional structural information is likely obtainable, irrespective of the fluorescence intensity contributions to the images. These results indicated the potential for clinical applications, such as depth diagnosis of early gastric cancer, for which quantitative depth estimation is currently difficult.

For example, in the clinical management of gastric cancer, the choice between endoscopic resection and surgical intervention is determined based on the depth of tumor invasion [23]. The accumulation of PpIX in early gastric cancer due to ALA administration has been reported in

several studies [24,25]. To facilitate minimally invasive treatments, Endoscopic Submucosal Dissection (ESD) is recommended for intramucosal cancers or those with submucosal invasion not exceeding 500 μm [26]. The average thickness of the gastric mucosa is approximately 1.3 mm [27]; therefore, it is imperative to accurately gauge the tumor's location within an approximate depth of 2 mm from the surface.

Compared to the numerical evaluations, both D and MAE exhibited poorer performance in the fluorescence experiments. One possible reason for this decrease may lie in the disparities between the simulated images used for training and the actual fluorescent images obtained during experiments. Figures 4(ii, iii) and 8(ii, iii) show the simulated and experimentally-acquired fluorescence images, respectively. The fluorescence distribution appears more expansive in the simulated images compared to its more localized spread in the experimental captures. This divergence could be ascribed to factors such as the limited detection sensitivity of the camera used in the experimental setup and quantization errors affecting intensity levels. The discrepancies between simulated and actual fluorescence images may contribute to the observed degradation in the network's estimation performance.

According to the mathematical model evaluation, when the optical property value increased by +15%, the MAE was 0.32 mm. The optical property values of the optical phantom employed in the experiment had a coefficient of variation of approximately 10% in the excitation wavelengths that significantly affected the estimation, suggesting that the estimation error in the experiment was influenced by the optical property values. In training the network, the optical property parameters were assumed constant. However, actual biological tissues have heterogeneity in optical properties [28], and the difference between assumed and actual optical property values becomes larger, which is expected to reduce the estimation accuracy. This problem can be addressed by creating a training dataset based on the expected range of optical property values for biological tissues and training the model. To construct a training dataset that reflects diverse optical properties, it is essential to perform a comprehensive evaluation of the optical properties of the targeted tissue. This will allow for the creation of a model that accounts for individual variations. It should be noted that the nature of these variations can differ substantially depending on the tissue type [29].

In the depth estimation, the target objects were limited to disc-shaped without thickness, and two PpIX pellets were arranged within the image. However, tumors have three-dimensional shapes, such as thickness, and the number of fluorescence objects included within the image is indefinite. For the three-dimensional shape of fluorescent objects, previous studies have shown that the distribution of thick fluorophores can be estimated using the linearity of fluorescence intensity ratios [30]. Similar to this study, a distribution might be estimated from fluorescence intensity ratio images in a deep learning model. Another limitation of this study is the simplified representation of PpIX concentrations. Both simulated and experimental fluorescence images used simplified PpIX concentrations to binary values. The actual concentrations of PpIX within tumor tissues are not binary by nature [3]. Consequently, future research targeting clinical applications would require the use of simulated fluorescence images that more accurately reflect the realistic PpIX distribution for training the networks.

In the adopted optical system, the irradiation position and angle were set such that the excitation light intensity distribution on the sample surface approached uniformity. In addition, the sample surface was horizontal and smooth. However, during actual observation of biological tissues, the endoscope tip angle is not constant with respect to the tissue, and it may be difficult to observe from the front, depending on the tumor's position. Therefore, the excitation light intensity distribution on the tissue and fluorophore is not uniform, and the reflection effect is significant [23]. To apply this model clinically, we assume real-time estimation of the depth of penetration through the endoscopic observation system. It is essential to utilize images that reproduce the

excitation light irradiation angle and intensity distribution within the camera field of view that may occur during endoscopic observation in the training dataset.

The PpIX pellet, localized within the optical phantom, was employed in the validations. Because adenocarcinoma, which accounts for 90% of gastric cancers, has the characteristic of invading deep into the tissue from the mucosal surface, it is necessary to estimate the position of the deepest part of the thick fluorophore to apply this model. Tumor depth information is crucial for determining the eligibility for ESD. However, in the case of gastric cancer, accurate depth diagnosis under endoscopy is currently difficult, for example, early gastric cancer sometimes is localized within the tissue without exposing the tumor to the mucosal surface [31]. In addition, cases have been reported where the diagnosis is affected by the appearance of low-grade dysplastic epithelium on the tumor surface, even after Helicobacter pylori eradication by conventional white-light imaging with magnifying videoendoscope [32]. The proposed method may be beneficial for diagnosing lesions localized within the mucosa, since the feasibility of implementing ESD eligibility diagnosis based on depth estimation under endoscopy is anticipated.

5. Conclusion

We proposed a method for estimating the three-dimensional distribution of fluorescent molecules within tissues using a deep learning model that utilizes the fluorescence intensity ratio images of two wavelengths observed on the tissue surface. We confirmed that the proposed method could estimate fluorescent molecules' planar position and depth. The estimation accuracy experimentally achieved a Dice coefficient of 0.73 for planar position estimation and an mean absolute error of 0.50 mm for depth estimation. This study does not reflect the optical characteristics of biological tissues and assumes that fluorescent molecules are inside the tissues. By creating a training dataset that matches the characteristics of the lesions targeted, such as gastric cancer, and training the model, we expect to improve the diagnostic accuracy in clinical applications.

Disclosures. The authors declare no conflicts of interest related to this article.

Data availability. Data underlying the results presented here are not publicly available but may be obtained from the authors upon reasonable request.

Supplemental document. See Supplement 1 for supporting content.

References

1. Q. T. Nguyen and R. Y. Tsien, "Fluorescence-guided surgery with live molecular navigation—a new cutting edge," Nat. Rev. Cancer **13**(9), 653–662 (2013).
2. P. Bou-Samra, N. Muhammad, and A. Chang, *et al.*, "Intraoperative molecular imaging: 3rd biennial clinical trials update," J. Biomed. Opt. **28**(05), 050901 (2023).
3. S. Kaneko and S. Kaneko, "Fluorescence-guided resection of malignant glioma with 5-ala," Int. J. Biomed. Imaging **2016**, 1–11 (2016).
4. H. Kostron, "Photodynamic diagnosis and therapy and the brain," Photodyn. Ther. Methods Protoc. **635**, 261–280 (2010).
5. D. Jocham, H. Stepp, and R. Waidelich, "Photodynamic diagnosis in urology: state-of-the-art," Eur. Urol. **53**(6), 1138–1150 (2008).
6. B. Krammer and K. Plaetzer, "Ala and its clinical impact, from bench to bedside," Photochem. Photobiol. Sci. **7**(3), 283–289 (2008).
7. Y. Harada, Y. Murayama, T. Takamatsu, E. Otsuji, and H. Tanaka, "5-aminolevulinic acid-induced protoporphyrin ix fluorescence imaging for tumor detection: Recent advances and challenges," Int. J. Mol. Sci. **23**(12), 6478 (2022).
8. H. Imanishi, T. Nishimura, and K. Awazu, "Depth estimation of protoporphyrin ix objects in turbid media considering the fluorescence intensity ratio between two wavelengths of light for application in invasion diagnosis of gastric cancer," Opt. Rev. **29**(4), 310–319 (2022).
9. K. K. Kolste, S. C. Kanick, P. A. Valdés, M. Jermyn, B. C. Wilson, D. W. Roberts, K. D. Paulsen, and F. Leblond, "Macroscopic optical imaging technique for wide-field estimation of fluorescence depth in optically turbid media for application in brain tumor surgical guidance," J. Biomed. Opt. **20**(2), 026002 (2015).

10. D. Wirth, K. Kolste, S. Kanick, D. W. Roberts, F. Leblond, and K. D. Paulsen, "Fluorescence depth estimation from wide-field optical imaging data for guiding brain tumor resection: a multi-inclusion phantom study," Biomed. Opt. Express **8**(8), 3656–3670 (2017).

11. C. M. O'Brien, K. W. Bishop, and H. Zhang, *et al.*, "Quantitative tumor depth determination using dual wavelength excitation fluorescence," Biomed. Opt. Express **13**(11), 5628–5642 (2022).

12. K. Kubota, J. Kuroda, M. Yoshida, K. Ohta, and M. Kitajima, "Medical image analysis: computer-aided diagnosis of gastric cancer invasion on endoscopic images," Surg. Endosc. **26**(5), 1485–1489 (2012).

13. S. Nagao, Y. Tsuji, and Y. Sakaguchi, *et al.*, "Highly accurate artificial intelligence systems to predict the invasion depth of gastric cancer: efficacy of conventional white-light imaging, nonmagnifying narrow-band imaging, and indigo-carmine dye contrast imaging," Gastrointest. Endoscopy **92**(4), 866–873.e1 (2020).

14. D. Ihara, H. Hazama, T. Nishimura, Y. Morita, and K. Awazu, "Fluorescence detection of deep intramucosal cancer excited by green light for photodynamic diagnosis using protoporphyrin ix induced by 5-aminolevulinic acid: an ex vivo study," J. Biomed. Opt. **25**(06), 1 (2020).

15. L. Wei, D. W. Roberts, N. Sanai, and J. T. Liu, "Visualization technologies for 5-ala-based fluorescence-guided surgeries," J. Neuro-Oncol. **141**(3), 495–505 (2019).

16. A. Kim, M. Roy, F. N. Dadani, and B. C. Wilson, "Topographic mapping of subsurface fluorescent structures in tissue using multiwavelength excitation," J. Biomed. Opt. **15**(6), 066026 (2010).

17. O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*, (Springer, 2015), pp. 234–241.

18. A. F. Agarap, "Deep learning using rectified linear units (relu)," arXiv, arXiv:1803.08375 (2018).

19. D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv, arXiv:1412.6980 (2014).

20. Q. Fang and D. A. Boas, "Monte carlo simulation of photon migration in 3d turbid media accelerated by graphics processing units," Opt. Express **17**(22), 20178–20190 (2009).

21. S. J. Ogbonna, W. Y. York, T. Nishimura, H. Hazama, H. Fukuhara, K. Inoue, and K. Awazu, "Increased fluorescence observation intensity during the photodynamic diagnosis of deeply located tumors by fluorescence photoswitching of protoporphyrin IX," J. Biomed. Opt. **28**(05), 055001 (2023).

22. Y. Shimojo, T. Nishimura, H. Hazama, T. Ozawa, and K. Awazu, "Measurement of absorption and reduced scattering coefficients in asian human epidermis, dermis, and subcutaneous fat tissues in the 400-to 1100-nm wavelength range for optical penetration depth and energy deposition analysis," J. Biomed. Opt. **25**(04), 1 (2020).

23. Y. Zhu, Q.-C. Wang, and M.-D. Xu, *et al.*, "Application of convolutional neural network in the diagnosis of the invasion depth of gastric cancer based on conventional endoscopy," World J. Gastroenterol. WJG **89**(4), 806–815.e1 (2019).

24. T. Namikawa, T. Yatabe, K. Inoue, T. Shuin, and K. Hanazaki, "Clinical applications of 5-aminolevulinic acid-mediated fluorescence for gastric cancer," World J. Gastroenterol. WJG **21**(29), 8769 (2015).

25. N. Koizumi, Y. Harada, T. Minamikawa, H. Tanaka, E. Otsuji, and T. Takamatsu, "Recent advances in photodynamic diagnosis of gastric cancer using 5-aminolevulinic acid," World J. Gastroenterol. **22**(3), 1289 (2016).

26. N. Uedo, Y. Takeuchi, and R. Ishihara, "Endoscopic management of early gastric cancer: endoscopic mucosal resection or endoscopic submucosal dissection: data from a japanese high-volume center and literature review," Annals of Gastroenterology **25**(4), 281 (2012).

27. H. Fujishima, T. Misawa, Y. Chijiwa, A. Maruoka, K. Akahoshi, and H. Nawata, "Scirrhous carcinoma of the stomach versus hypertrophic gastritis: findings at endoscopic us," Radiology **181**(1), 197–200 (1991).

28. J. Schmitt and G. Kumar, "Turbulent nature of refractive-index variations in biological tissue," Opt. Lett. **21**(16), 1310–1312 (1996).

29. S. L. Jacques, "Optical properties of biological tissues: a review," Phys. Med. Biol. **58**(11), R37–R61 (2013).

30. M. Kirillin, A. Khilov, and D. Kurakina, *et al.*, "Dual-wavelength fluorescence monitoring of photodynamic therapy: from analytical models to clinical studies," Cancers **13**(22), 5807 (2021).

31. Y. Yamamoto, J. Fujisaki, and T. Hirasawa, *et al.*, "Therapeutic outcomes of endoscopic submucosal dissection of undifferentiated-type intramucosal gastric cancer without ulceration and preoperatively diagnosed as 20 millimetres or less in diameter," Dig. Endosc. **22**(2), 112–118 (2010).

32. A. Saka, K. Yagi, and S. Nimura, "Endoscopic and histological features of gastric cancers after successful helicobacter pylori eradication therapy," Gastric Cancer **19**(2), 524–530 (2016).